1861

LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

First draft--comments welcomed

Copy 1

DEWEY

EXTENDED NATURAL CONJUGATE DISTRIBUTIONS FOR THE MULTINORMAL PROCESS

Kinosuke

Albert Ando and G. M. Kaufman

80-64

HD28
.M414
[illegible]

0. Introduction

In [1] we developed the distribution theory necessary to carry out Bayesian analysis of the multivariate Normal process when neither the mean vector nor the variance-covariance matrix of the process is known with certainty. The class of natural conjugate distributions of the process was identified and defined in formula (6) of [1]. We labelled members of this class "Normal-Wishart" and used Normal-Wishart priors to carry out prior to posterior and pre-posterior analysis. In section 1.3 of [1] we pointed out that the Normal-Wishart class lacks flexibility in a certain sense.

Furthermore, Rothenburg [4] points out an additional difficulty associated with natural conjugate analysis of the Multinormal process when this analysis is applied to reduced form systems; namely, if the joint prior of $(\tilde{\underline{\underline{\Pi}}}, \tilde{\underline{\underline{h}}})$, the matrix of coefficients of the system and the inverse of the variance-covariance matrix of the residual error vector respectively, is in the natural conjugate family, then the marginal variances of all coefficients in the ith row of $\tilde{\underline{\underline{\Pi}}}$ are proportional to the marginal variances of coefficients in the jth row of $\tilde{\underline{\underline{\Pi}}}$. In practical problems, such a restriction is often undesirable. Zellner [5] attempts to deal with this problem.

Martin [3] highlights another restrictive feature of the Normal-Wishart family of priors. The parameter $\underline{\underline{V}}$ of a Normal-Wishart prior as defined in (2) below is $(\nu-2)$ times the expected value $E(\tilde{\underline{\underline{h}}}^{-1})$ of the variance-covariance matrix of the data generating process (1). He goes on to show that given $(\underline{\underline{V}}, \underline{\nu})$ and letting $h^{-1}_{\alpha\beta}$ be the (α, β) element of the $(r \times r)$ matrix $\tilde{\underline{\underline{h}}}^{-1}$, for $1 \le \alpha, \beta, \gamma, \delta \le r$,

$$\mathrm{Var}(\tilde{h}^{-1}_{\alpha\beta}) = \frac{1}{3(\nu-2)(\nu-4)} \left[\frac{8-\nu}{\nu-2} V^2_{\alpha\beta} + V_{\alpha\alpha} V_{\beta\beta}\right]$$

and

$$\mathrm{Cov}(\tilde{h}^{-1}_{\alpha\beta}, \tilde{h}^{-1}_{\gamma\delta}) = \frac{1}{3(\nu-2)(\nu-4)} \left[\frac{2(5-\nu)}{\nu-2} V_{\alpha\beta}V_{\gamma\delta} + V_{\alpha\gamma}V_{\beta\delta} + V_{\alpha\delta}V_{\beta\gamma}\right] .$$

Thus specification of ν and $\underline{\underline{V}}$ *completely specifies the variance of all elements of the variance-covariance matrix of the process and fixes the covariances among elements of this matrix as well*. This is generally an undesirable state of affairs as one may reasonably wish to set the mean of $\underline{\underline{\tilde{h}}}^{-1}$ *and* the variances and covariances among its elements in a way that does not jive with the above formulae. In

particular, note that the (prior) variance of the variances of the process are, for $\alpha=1,2,\ldots,r$, restricted to be

$$\mathrm{Var}(\tilde{h}^{-1}_{\alpha\alpha}) = \frac{2v^2_{\alpha\alpha}}{(\nu-2)^2(\nu-4)} = \frac{2}{\nu-4}\,E^2(\tilde{h}^{-1}_{\alpha\alpha})$$

Our purpose here is to extend the Normal-Wishart class of priors so as to allow much greater flexibility in the assignment of priors to the parameter of a multivariate Normal data generating process, while retaining as much analytical simplicity as possible. In particular, we identify an extension that liberates us from the tyranny outlined in the preceding paragraph and allows us to assign (prior) variances much more freely to the diagonal elements of $\underline{\underline{\tilde{h}}}^{-1}$--the marginal variances of elements of $\underline{\tilde{\mu}}$, the mean vector of the data generating process.

Throughout we use the notation and conventions adopted in [1].

1.1 Natural Conjugate Distributions of $(\underline{\tilde{\mu}}, \underline{\underline{\tilde{h}}})$, $\underline{\tilde{\mu}}$, and $\underline{\underline{\tilde{h}}}$

We define an r-dimensional Independent Multinormal process as one that generates independent r x 1 random vectors $\underline{\tilde{x}}^{(1)},\ldots,\underline{\tilde{x}}^{(j)},\ldots$ with identical densities

$$f^{(r)}_N(\underline{x}|\underline{\mu}, \underline{\underline{h}}) = (2\pi)^{-\frac{1}{2}r}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^t\underline{\underline{h}}(\underline{x}-\underline{\mu})}\, |\underline{\underline{h}}|^{\frac{1}{2}} \qquad (1)$$

$$-\underline{\infty} < \underline{x} < +\underline{\infty}\,,$$
$$-\underline{\infty} < \underline{\mu} < +\underline{\infty}\,,$$
$$\underline{\underline{h}} \text{ is PDS}\quad .$$

When both $\underline{\tilde{\mu}}$ and $\underline{\underline{\tilde{h}}}$ are random variables, the natural conjugate of (1) is the Normal-Wishart distribution $f^{(r)}_{NW}(\underline{\mu}, \underline{\underline{h}}|\underline{m}, \underline{\underline{V}}, n, \nu)$ defined as equal to

$$k(r,\nu)\, e^{-\frac{1}{2}n(\underline{\mu}-\underline{m})^t \underline{\underline{h}}(\underline{\mu}-\underline{m})}\, |\underline{\underline{h}}|^{\frac{1}{2}\delta}\, e^{-\frac{1}{2}\mathrm{tr}\, \underline{\underline{h}}\, \underline{\underline{V}}^*}\, |\underline{\underline{V}}^*|^{\frac{1}{2}(\nu+r-1)}\, |\underline{\underline{h}}|^{\frac{1}{2}\nu-1}$$

$$= \begin{cases} f_N^{(r)}(\underline{\mu}|\underline{m}, \underline{\underline{h}}n)\; f_W^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \nu) & \text{if } n \geq 0,\ \nu > 0\ , \\ 0 & \text{otherwise} \end{cases} \qquad (2a)$$

where

$$\underline{\underline{V}}^* = \begin{cases} \underline{\underline{V}} & \text{if } \underline{\underline{V}} \text{ is PDS} \\ \underline{\underline{0}} & \text{otherwise} \end{cases} \qquad (2b)$$

$$\delta = \begin{cases} 1 & \text{if } n > 0 \\ 0 & \text{if } n = 0 \end{cases} \qquad (2c)$$

and

$$k(r,\nu) = 2^{-\frac{1}{2}r(\nu+r)}\, \pi^{-\frac{1}{4}r(r+1)}\, n^{\frac{1}{2}r\delta}\, [\prod_{i=1}^{r} \Gamma(\tfrac{1}{2}(\nu+r-i))]^{-1}\ . \qquad (2d)$$

2. Some Extensions

Many possibilities for generalizing (2) exist. We shall restrict our attention here to generalizations which (i) are closed under the binary operation of going from a prior to a posterior distribution of $(\underline{\mu}, \underline{\underline{h}})$ in that both prior and posterior distributions have the same functional form, (ii) allow us to easily derive marginal distributions of $\underline{\mu}$ and of $\underline{\underline{h}}$ and (iii) enable us to do preposterior analysis in a simple fashion.

The extensions exploited here use generalizations of the Wishart distribution due to Bellman [2].

To foreshadow the effect of extending the natural conjugate family, we tabulate some moments. Let

$$\underline{\underline{h}}^{-1} = \underline{\underline{\Sigma}} = [\sigma_{ij}]\ , \quad \dim(r \times r)$$

$$\underline{h}^{-1} = \underline{\sigma} = (\sigma_{11},\ldots,\sigma_{1r},\sigma_{22},\ldots,\sigma_{rr})^t\ , \quad \dim(1 \times r(r+1))\ ,$$

$$\underline{\bar{h}}^{-1} = \underline{\bar{\sigma}} = E(\underline{\tilde{\sigma}})\ ,$$

$$V(\tilde{\underline{\sigma}}) = E(\tilde{\underline{\sigma}}-\bar{\underline{\sigma}})(\tilde{\underline{\sigma}}-\bar{\underline{\sigma}})^t \quad ,$$

$$V\begin{pmatrix}\tilde{\underline{\mu}}\\ \tilde{\underline{\sigma}}\end{pmatrix} = \begin{bmatrix} V(\tilde{\underline{\mu}}) & E(\tilde{\underline{\mu}}-\underline{m})(\tilde{\underline{\sigma}}-\bar{\underline{\sigma}})^t \\ \left[E(\tilde{\underline{\mu}}-\underline{m})(\tilde{\underline{\sigma}}-\bar{\underline{\sigma}})^t\right]^t & V(\tilde{\underline{\sigma}}) \end{bmatrix}$$

In Table I we compare $E(\tilde{\underline{\mu}})$, $\bar{\underline{\sigma}}$, $V(\tilde{\underline{\sigma}})$, and $V\begin{pmatrix}\tilde{\underline{\mu}}\\ \tilde{\underline{\sigma}}\end{pmatrix}$ when we assign a Normal-Wishart prior with parameter $(\underline{m}, n, \underline{\underline{V}}, \nu)$ to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ with these same moments when an extended Type 1 Normal-Wishart prior with parameter $(\underline{m}, n, \underline{\underline{V}}, \underline{\nu})$, $\underline{\nu} \equiv (\nu_1,\ldots,\nu_r)$ is assigned to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$. Here $\underline{\underline{T}}$ is an $(r \times r)$ upper triangular matrix such that $\underline{\underline{T}}^t\underline{\underline{T}} = \underline{\underline{V}}^{-1}$, $\underline{\eta}$ is an $(r \times r)$ diagonal matrix whose diagonal elements are functions of $\underline{\nu}$ alone and are later shown to fulfill the recursion relation[†]

$$\eta_{jj} = \frac{a_{r-j+2}-\frac{1}{2}(r-j+2)}{a_{r-j+1}-\frac{1}{2}(r-j+2)}\,\eta_{j-1,j-1} \;, \qquad j=2,3,\ldots,r \;.$$

Elements of $\underline{\underline{W}}(\underline{\underline{V}}, \nu)$ below are shown by Martin to be

$$W_{ij}(\underline{\underline{V}}, \nu) = \mathrm{cov}(\tilde{\sigma}_{\alpha\beta}, \tilde{\sigma}_{\gamma\delta}) = \frac{1}{3(\nu-2)(\nu-4)}\left[\frac{2(5-\nu)}{\nu-2}V_{\alpha\beta}V_{\gamma\delta}+V_{\alpha\gamma}V_{\beta\delta}+V_{\alpha\delta}V_{\beta\gamma}\right] \;.$$

$$W_{ii}(\underline{\underline{V}}, \nu) = \mathrm{var}(\tilde{\sigma}_{\alpha\beta}) = \frac{1}{3(\nu-2)(\nu-4)}\left[\frac{8-\nu}{\nu-2}V_{\alpha\beta}^2 + V_{\alpha\alpha}V_{\beta\beta}\right] \;. \qquad i,j=1,2,\ldots,r,$$

and we show that when $r=2$,[‡]

$$\psi_{11}(\underline{\underline{T}}, \underline{\nu}) = \mathrm{var}(\tilde{\sigma}_{11}) = \frac{E^2(\tilde{\sigma}_{11})}{2(a_2-\frac{5}{2})}$$

and

$$\psi_{33}(\underline{\underline{T}}, \underline{\nu}) = \frac{1}{4(a_2-\frac{3}{2})(a_2-\frac{5}{2})}\left\{(t_{21}^{-1})^4+2(t_{21}^{-1})^2(t_{22}^{-1})^2\frac{\eta_{22}}{\eta_{11}}+(t_{22}^{-1})^4\frac{\eta_{22}}{\eta_{11}}\cdot\frac{a_2-2}{a_1-1}\right\} - E^2(\tilde{\sigma}_{22}) \;,$$

[†] The a_is are functions of $\underline{\nu}$ alone, as defined in (4c) below.

[‡] The extension of the results shown here to arbitrary r is straight forward, but tedious, and will be presented in a later draft.

where t^{-1}_{ij} denotes the ijth element of $\underline{\underline{T}}^{-1}$.

TABLE 1

		Normal-Wishart†	Extended Type 1 Normal-Wishart
$E(\underline{\tilde{\mu}})$	:	$\underline{m}'$	$\underline{m}'$
$E(\underline{\underline{\tilde{\sigma}}}) = V(\underline{\tilde{\mu}})$	:	$\frac{1}{\nu-2}\,\underline{\underline{V}}$	$\underline{\underline{T}}^{-1}\,\underline{\underline{\eta}}\,\underline{\underline{T}}^{t^{-1}}$
$V(\underline{\tilde{\sigma}})$	:	$\underline{\underline{W}}(\underline{\underline{V}},\ \nu)$	$\underline{\underline{\psi}}(\underline{\underline{T}},\ \underline{\nu})$
$V\begin{pmatrix}\underline{\tilde{\mu}}\\ \underline{\tilde{\sigma}}\end{pmatrix}$	:	$\begin{bmatrix}\frac{1}{\nu-2}\,\underline{\underline{V}} & \underline{\underline{0}}\\ \underline{\underline{0}} & \underline{\underline{W}}(\underline{\underline{V}},\ \nu)\end{bmatrix}$	$\begin{bmatrix}\underline{\underline{T}}^{-1}\,\underline{\underline{\eta}}\,\underline{\underline{T}}^{t^{-1}} & \underline{\underline{0}}\\ \underline{\underline{0}} & \underline{\underline{\psi}}(\underline{\underline{T}},\ \underline{\nu})\end{bmatrix}$

In a later section we discuss a numerical example and compare the above.

2.1 Bellman's Extension of an Integral Identity of Siegel

In [2] Bellman generalizes Siegel's identity:

$$\Gamma_r(\nu)\ |\underline{\underline{V}}|^{-\frac{1}{2}(\nu+r-1)} = \int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}}\ |\underline{\underline{h}}|^{\frac{1}{2}\nu-1} d\underline{\underline{h}}\ , \tag{3}$$

$\underline{\underline{V}}$ is (r x r) PDS ,

$R_{\underline{\underline{h}}} \equiv \{\underline{\underline{h}} \mid \underline{\underline{h}}$ is (r x r) PDS$\}$,

$\nu > 0$.

---

†Martin [3].

and

$$\Gamma_r(\nu) \equiv \pi^{\frac{1}{4}r(r-1)} \; \Gamma(\nu)\; \Gamma(\nu+\tfrac{1}{2}) \ldots \Gamma(\nu+\tfrac{1}{2}(r-1)) \quad .$$

Letting $\underline{\underline{h}}$ be an $(r \times r)$ real symmetric matrix, $\underline{\underline{V}}^{[\ell]}=(\underline{\underline{V}}_{ij})$, $1 \leq i,\ j \leq \ell$, and $\underline{\underline{h}}_{[\ell]}=(h_{ij})$, $\ell \leq i,\ j \leq r$, he shows in effect that

$$\int_{R_{\underline{\underline{h}}}} \frac{e^{-\frac{1}{2}\mathrm{tr}\, \underline{\underline{h}}\, \underline{\underline{V}}} \; |\underline{\underline{h}}|^{\frac{1}{2}\nu_r - 1}}{\prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{\nu_{j-1}}} \, d\underline{\underline{h}} = 2^{\alpha} \, \pi^{\frac{1}{4}r(r-1)} \Gamma_r(\underline{a}) \; |\underline{\underline{V}}|^{\beta - \frac{1}{2}(\nu_r + r - 1)} \prod_{j=1}^{r-1} |\underline{\underline{V}}^{[j]}|^{-\nu_j} \tag{4a}$$

where

$$\alpha = \tfrac{1}{2}r(\nu_r + r - 1) - \sum_{j=2}^{r} (r-j+1)\nu_{j-1} \quad , \quad \beta = \sum_{j=1}^{r-1} \nu_j \quad , \tag{4b}$$

$$a_r = \tfrac{1}{2}(\nu_r + r - 1) \quad , \quad a_j = a_r - \sum_{k=1}^{r-j} \nu_k \quad , \quad j=1,2,\ldots,r-1 \; , \tag{4c}$$

$$\underline{a} = (a_1, a_2, \ldots, a_r) \quad , \quad \Gamma_r(\underline{a}) = \prod_{j=1}^{r} \Gamma(a_j - \tfrac{1}{2}(j-1)) \quad . \tag{4d}$$

and the ν_j must be such that $a_j > \frac{1}{2}(j-1)$. We shall call the integrand in (4) the kernel of a Type 1 extended Wishart density $f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu})$, with parameter $(\underline{\underline{V}}, \underline{\nu})$ where $\underline{\nu} \equiv (\nu_1, \ldots, \nu_r)$. Henceforth we shall let

$$\Gamma_r^*(\underline{\nu}) \equiv 2^{\alpha} \, \pi^{\frac{1}{4}r(r-1)} \; \Gamma(a_r - \tfrac{1}{2}(r-1)) \ldots \Gamma(a_2 - \tfrac{1}{2}) \; \Gamma(a_1)$$

for notational simplicity.

This generalization of the Wishart distribution is particularly well suited for our purposes, as the peculiarly simple form in which principal minors of $\underline{\underline{V}}$ appear allow us to do preposterior as well as prior to posterior analysis in a very simple fashion.

2.2 Extended Natural Conjugate Distributions of $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$

Using (4) allows us to define an extended natural conjugate distribution of $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ with $(r-1)$ more parameters than the Normal-Wishart density when the data generating process is $(r \times 1)$, giving a great deal more flexibility in the assignment of priors to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$.

We define a type 1 extended Normal-Wishart distribution $f_{NW1}^{(r)}(\underline{\mu},\underline{\underline{h}}|\underline{m},\underline{\underline{V}},n,\underline{\nu})$ as equal to

$$f_N^{(r)}(\underline{\mu}|\underline{m},\ \underline{\underline{h}}\ n)\ f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}},\ \underline{\nu})$$
$$= 2^{-\frac{1}{2}r(\nu_r+r-1)}\left[\frac{\Gamma_r^*(\underline{\nu})}{\Gamma_r(\nu_r)}\right]^{-1}\prod_{j=1}^{r-1}|\underline{\underline{V}}^{[j]}|^{-\nu_j}\ \prod_{j=2}^{r}|\underline{\underline{h}}_{[j]}|^{-\nu_{j-1}}\ f_{NW}^{(r)}(\underline{\mu},\underline{\underline{h}}|\underline{m},\underline{\underline{V}},n,\nu_r) \tag{5}$$

If (5) is to be a proper density function $\underline{\underline{V}}$ must be PDS, $\nu_r > 0$, and $n > 0$. In addition, the ν_j must be small enough so that $a_j > \frac{1}{2}(j-1)$, $j=2,\ldots,r$. Otherwise, the ν_j are unrestricted in sign.

We obtain the marginal prior of $\tilde{\underline{\mu}}$ by integrating (5) with respect to $\underline{\underline{h}}$. If $a_j > \frac{1}{2}(j-1)$ for $j=1,2,\ldots,r$, $n > 0$, $\nu_r > 0$, and $\underline{\underline{V}}$ is PDS then

$$D(\underline{\mu}|\underline{m},\ \underline{\underline{V}},\ \underline{\nu}) \propto ((\nu_r-2\beta)+(\underline{m}-\underline{\mu})^t\underline{\underline{H}}_\nu(\underline{m}-\underline{\mu}))^{\beta-\frac{1}{2}(\nu_r+r)}$$
$$\cdot \prod_{j=1}^{r-1}(2\nu_j-j) + (\underline{m}-\underline{\mu})^{[j]^t}\ \underline{\underline{H}}_j(\underline{m}-\underline{\mu})^{[j]})^{-\nu_j} \tag{6a}$$

where $\underline{\underline{H}}_j \equiv (2\nu_j-j)n\ \underline{\underline{V}}^{[j]^{-1}}$ and $\underline{\underline{H}}_\nu = (\nu_r-2\beta)n\ \underline{\underline{V}}^{-1}$. When $\underline{\underline{H}}_\nu$ and all $\underline{\underline{H}}_j$, $j=1,2,\ldots,r-1$, are PDS, then

$$D(\underline{\mu}|\underline{m},\underline{\underline{V}},\nu) \propto f_S^{(r)}(\underline{\mu}|\underline{m},\underline{\underline{H}}_\nu,\nu_r-2\beta)\ \prod_{j=1}^{r-1} f_S^{(r)}(\underline{\mu}^{[j]}|\underline{m}^{[j]},\underline{\underline{H}}_j,2\nu_j-j) \tag{6b}$$

where $f_S^{(j)}(\underline{\mu}^{[j]}|\underline{m}^{[j]},\ \underline{\underline{H}}_j,\ 2\nu_j-j)$ is the non-degenerate multivariate Student distribution.

Proof: Integrating over $R_{\underline{\underline{h}}}$ we have

$$D(\underline{\mu}|\underline{m},\underline{\underline{V}},n,\underline{\nu})\ \int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\{n(\underline{\mu}-\underline{m})(\underline{\mu}-\underline{m})^t+\underline{\underline{V}}\}}\ |\underline{\underline{h}}|^{\frac{1}{2}(\nu_r+\delta)-1}\ \prod_{j=2}^{r}|\underline{\underline{h}}_{[j]}|^{-\nu_{j-1}}\ d\underline{\underline{h}}\quad .$$

The integrand above is the kernel of an extended Type 1 Wishart density with

parameter $(n(\underline{m}-\underline{\mu})(\underline{m}-\underline{\mu})^t+\underline{\underline{V}},\ \underline{\nu})$ where $\underline{\nu} = (\nu_1,\nu_2,\ldots,\nu_{r-1},\nu_r+\delta)$. Hence defining

$$\underline{\mu}^{[j]} = (\mu_1,\ldots,\mu_j)^t$$

we have by use of (4a),

$$D(\underline{\mu}|\underline{m},\underline{\underline{V}},n,\underline{\nu}) \propto |n(\underline{\mu}-\underline{m})(\underline{\mu}-\underline{m})^t+\underline{\underline{V}}|^{-a_1-\frac{1}{2}\delta} \cdot \prod_{j=1}^{r-1} |[n(\underline{\mu}-\underline{m})(\underline{\mu}-\underline{m})^t+\underline{\underline{V}}^{[j]}|^{-\nu_j}$$

$$= (1 + (\underline{\mu}-\underline{m})^t(n\underline{\underline{V}}^{-1})(\underline{\mu}-\underline{m}))^{-a_1-\frac{1}{2}\delta} \cdot \prod_{j=1}^{r-1} (1+(\underline{\mu}-\underline{m})^{[j]^t}(n\underline{\underline{V}}^{[j]^{-1}})(\underline{\mu}-\underline{m})^{[j]})^{-\nu_j}.$$

Provided that $\underline{\underline{V}}$ is PDS, $\nu_r-2\beta > 0$, $n > 0$, and $2\nu_j-j > 0$, $j=1,2,\ldots,r-1$, $\underline{\underline{H}}_j \equiv n(2\nu_j-j)\underline{\underline{V}}^{[j]^{-1}}$ is PDS, $j=1,2,\ldots,r-1$, $\underline{\underline{H}}_\nu \equiv (\nu_r-2\beta)n\underline{\underline{V}}^{-1}$ is PDS, $\delta=1$, and we have (6b). If $\underline{\underline{V}}$ is PDS and $n > 0$, (6a) holds even when one or more $2\nu_j-j \leq 0$ as long as $a_j > \frac{1}{2}(j-1)$, $j=1,2,\ldots,r$.

If one or more of the conditions $\nu_r > 0$, $\underline{\underline{V}}$ PDS, $a_j > \frac{1}{2}(j-1)$, $j=1,2,\ldots,r$ are violated, the marginal distribution of $\tilde{\underline{\mu}}$ does not exist.

The constant that normalizes (6a) is

$$S(\underline{\underline{V}},\underline{\nu}) = \kappa(\underline{\nu})\ 2^r|\underline{\underline{V}}|^{-\frac{1}{2}}\ \{\prod_{j=1}^{r} B(\tfrac{1}{2},\ a_j-\tfrac{1}{2}(j-1))\}^{-1} \tag{6c}$$

where

$$\kappa(\underline{\nu}) = (\nu_r-2\beta)^{-a_1-\frac{1}{2}} \prod_{j=1}^{r-1} (2\nu_j-j)^{\nu_j} \tag{6d}$$

Proof: Since $\underline{\underline{V}}$ is PDS, there exists an upper triangular matrix $\underline{\underline{T}}$ such that $\underline{\underline{T}}^t\underline{\underline{T}} = \underline{\underline{V}}^{-1}$. It is also true that $\underline{\underline{T}}^{[j]^t}\underline{\underline{T}}^{[j]}=\underline{\underline{V}}^{[j]^{-1}}$ for $j=1,2,\ldots,r-1$.

Let $\underline{y}^{[j]}=\underline{\underline{T}}^{[j]}(\underline{\mu}-\underline{m})^{[j]}$ and make the integrand transformation from $\underline{\mu}$ to $\underline{y}$ in (6a). Since the Jacobian of the transformation is $|\underline{\underline{V}}|^{\frac{1}{2}}$,

$$D(\underline{y}|\underline{\nu}) = |\underline{\underline{V}}|^{\frac{1}{2}}\ \kappa(\underline{\nu})\ (1+\underline{y}^t\underline{y})^{-a_1-\frac{1}{2}} \prod_{j=1}^{r-1} (1+\underline{y}^{[j]^t}\underline{y}^{[j]})^{-\nu_j}\ . \tag{6e}$$

Integrating over $R_{\underline{y}} \equiv \{\underline{y}|-\underline{\infty} < y < +\underline{\infty}\}$ yields the normalizing constant. First we evaluate

$$\int_{R_{\underline{y}}} (1 + \underline{y}^t\underline{y})^{-a_1-\frac{1}{2}} \prod_{j=1}^{r-1} (1 + \underline{y}^{[j]^t}\underline{y}^{[j]})^{-\nu_j}\ d\underline{y} \tag{6f}$$

$$= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (1 + \sum_{k=1}^{r} y_k^2)^{-a_1-\frac{1}{2}} \prod_{j=1}^{r-1} (1 + \sum_{\ell=1}^{j} y_\ell^2)^{-\nu_j} \, dy_1 \ldots dy_r \quad .$$

Make the integrand transform $y_k^2 = x_k$ in (6e) and write it as

$$2^{-r} \int_0^{\infty} \cdots \int_0^{\infty} (1 + \sum_{k=1}^{r} x_k)^{-a_1-\frac{1}{2}} [\prod_{j=1}^{r-1} (1 + \sum_{\ell=1}^{j} x_\ell)^{-\nu_j}][\prod_{i=1}^{r} x_i^{-\frac{1}{2}}] \, dx_1 \ldots dx_r \quad . \tag{6g}$$

Integrating first with respect to x_r, we find that the above integral equals

$$2^{-r} B(\tfrac{1}{2}, a_1) \int_0^{\infty} \cdots \int_0^{\infty} (1 + \sum_{k=1}^{r-1} x_k)^{-a_2} [\prod_{j=1}^{r-2} (1 + \sum_{\ell=1}^{j} x_\ell)^{-\nu_j}][\prod_{i=1}^{r-1} x_i^{-\frac{1}{2}}] \, dx_1 \ldots dx_{r-1} \quad .$$

Continuing in a similar fashion yields (6f) as equal to

$$2^{-r} B(\tfrac{1}{2}, a_1) \, B(\tfrac{1}{2}, a_2 - \tfrac{1}{2}) \ldots B(\tfrac{1}{2}, a_j - \tfrac{1}{2}(j-1)) \ldots B(\tfrac{1}{2}, a_r - \tfrac{1}{2}(r-1)) \quad ,$$

and (6c) follows directly.

Later we will need the following

Theorem 1: Let $\tilde{\underline{x}}^{(1)}, \ldots, \tilde{\underline{x}}^{(j)}, \ldots, \tilde{\underline{x}}^{(n)}$ be $(r \times 1)$ mutually independent random vectors which, conditional on $\tilde{\underline{\underline{h}}} = \underline{\underline{h}}$, are distributed according to $f_N^{(r)}(\cdot | \underline{\bar{x}}, \underline{\underline{h}} n_j)$, $n_j > 0$, $j=1,2,\ldots,n$; let $\tilde{\underline{\underline{h}}}$ have an extended Type 1 Wishart distribution $f_{W1}^{(r)}(\underline{\underline{h}} | \underline{\underline{V}}, \underline{\nu})$ and define $\tilde{\underline{y}}$ as an $(r \times 1)$ linear combination of the $\tilde{\underline{x}}$s,

$$\tilde{\underline{y}} = \underline{\alpha} + \Sigma \, \beta_j \, \tilde{\underline{x}}^{(j)} \quad ,$$

where β_j's are constants. Then, provided that $\nu_r > 0$, $a_j > \frac{1}{2}(j-1)$, $j=1,2,\ldots,r$, and $\underline{\underline{V}}$ is PDS, the distribution of $\tilde{\underline{y}}$ is of the form (6a)

$$D(\underline{y}) \propto (1 + (\underline{y} - \underline{\bar{y}})^t (k\underline{\underline{V}}^{-1})(\underline{y} - \underline{\bar{y}}))^{\beta - \frac{1}{2}(\nu_r + r)}$$

$$\cdot \prod_{j=1}^{r-1} (1 + (\underline{y} - \underline{\bar{y}})^{[j]^t} (k\underline{\underline{V}}^{[j]^{-1}})(\underline{y} - \underline{\bar{y}})^{[j]})^{-\nu_j}$$

where

$$k = (\Sigma \frac{\beta_j^2}{n_j})^{-1} \quad .$$

If, in addition $\nu_r - 2\beta > 0$ and $2\nu_j - j \geq 0$,

$$D(\underline{y}) \propto f_S^{(r)}(\underline{y}|\bar{\underline{y}}, k\underline{\underline{V}}^{-1}, \nu_r-2\beta) \prod_{j=1}^{r-1} f_S^{(j)}(\underline{y}^{[j]}|\bar{\underline{y}}^{[j]}, k\underline{\underline{V}}^{[j]^{-1}}, 2\nu_j-j) \quad .$$

Proof: Since, conditional on $\tilde{\underline{\underline{h}}}=\underline{\underline{h}}$, the $\tilde{\underline{x}}$ s are each independent (r x 1) Normal vectors,

$$\tilde{\underline{y}}|\underline{\underline{h}} \sim f_N^{(r)}(\underline{y}|\bar{\underline{y}}, k\underline{\underline{h}})$$

where

$$\bar{\underline{y}} = \underline{\alpha} + \Sigma\, \beta_j\, \bar{\underline{x}}^{(j)} \quad \text{and} \quad k = \left(\Sigma \frac{\beta_j^2}{n_j}\right)^{-1} \quad .$$

Integrating with respect to $\underline{\underline{h}}$ gives

$$D(y) = \int_{R_{\underline{\underline{h}}}} f_N^r(\underline{y}|\bar{\underline{y}}, \underline{\underline{h}}k)\ f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu})d\underline{\underline{h}}$$

$$\propto \int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\{k(\underline{y}-\bar{\underline{y}})(\underline{y}-\bar{\underline{y}})^t+\underline{\underline{V}}\}}\ |\underline{\underline{h}}|^{\frac{1}{2}(\nu_r+1)-1} \cdot \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu_{j-1}}\ d\underline{\underline{h}} \quad .$$

The integrand above is the kernel of an extended Type 1 Wishart density with parameter $(k(\underline{y}-\bar{\underline{y}})(\underline{y}-\bar{\underline{y}})^t+\underline{\underline{V}}, \underline{\nu}_o)$ where $\underline{\nu}_o=(\nu_1,\ldots,\nu_{r-1},\nu_r+1)$. Formula (7) follows directly.

The marginal prior on $\tilde{\underline{\underline{h}}}$ is obtained by integrating (5) with respect to $\tilde{\underline{\mu}}$. If $n \geq 0$, $\underline{\underline{V}}$ is PDS, and $a_j > \frac{1}{2}(j-1)$, then

$$D(\underline{\underline{h}}|\underline{m},\underline{\underline{V}},n,\underline{\nu}) = f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu}) \propto e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}}\ |\underline{\underline{h}}|^{\frac{1}{2}\nu_r-1} \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu_{j-1}} \qquad (7)$$

Proof: Integrate (5) over $-\underline{\infty} < \underline{\mu} < +\infty$:

$$D(\underline{\underline{h}}|\underline{m}, \underline{\underline{V}}, n, \underline{\nu}) = \int_{R_{\underline{\mu}}} f_N^{(r)}(\underline{\mu}|\underline{m}, \underline{\underline{h}}n)\ f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu})d\underline{\mu}$$

$$= f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu}) \int_{R_{\underline{\mu}}} f_N^{(r)}(\underline{\mu}|\underline{m}, \underline{\underline{h}}n)d\underline{\mu} = f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}, \underline{\nu}) \quad .$$

If a type 1 extended Normal-Wishart distribution with parameter $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}')$ is assigned to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ and if a sample then yields a statistic

$(\underline{m}, \underline{\underline{V}}, n, \nu_r)$ the posterior distribution of $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ will be type 1 extended Normal-Wishart with parameter $(\underline{m}'', \underline{\underline{V}}^{*''}, n'', \underline{\nu}'')$ where

$$n'' = n'+n \quad , \quad \delta'' = \begin{cases} 1 & \text{if} \quad n'' > 0 \\ 0 & \phantom{\text{if}} \quad n'' = 0 \end{cases} \quad , \quad \underline{m}'' = n''^{-1}(n'\underline{m}'+n\,\underline{m}) \ , \tag{8a}$$

$$\begin{aligned} &\nu''_r = \nu'_r + \nu + r + \delta + \delta' - \delta'' - \Phi - 1 \quad , \quad \nu''_j = \nu'_j \quad , \quad 1 \le j \le r-1 \quad , \\ &\underline{\nu}'' = (\nu''_1, \ldots, \nu''_r) \quad , \end{aligned} \tag{8b}$$

$$\underline{\underline{V}}^{*''} = \begin{cases} \underline{\underline{V}}' + \underline{\underline{V}} + n'\underline{m}'\underline{m}'^t + n\,\underline{m}\,\underline{m}^t - n''\underline{m}''\underline{m}''^t \equiv \underline{\underline{V}}'' & \text{if } \underline{\underline{V}}'' \text{ is PDS} \\ \underline{\underline{0}} & \text{otherwise} \end{cases} \quad . \tag{8c}$$

Proof: As in the proof of (7) in [1], we multiply the kernel of the prior density by the kernel of the likelihood, obtaining

$$\begin{aligned} &e^{-\frac{1}{2}n'(\underline{\mu}-\underline{m}')^t\underline{\underline{h}}(\underline{\mu}-\underline{m}')} \ |\underline{\underline{h}}|^{\frac{1}{2}\delta'} \ e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}'} \ |\underline{\underline{h}}|^{\frac{1}{2}\nu'_r-1} \ \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu'_{j-1}} \\ &\cdot e^{-\frac{1}{2}n(\underline{m}-\underline{\mu})^t\underline{\underline{h}}(\underline{m}-\underline{\mu})} \ |\underline{\underline{h}}|^{\frac{1}{2}\delta} \ e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}} \ |\underline{\underline{h}}|^{\frac{1}{2}(\nu+r-\Phi-1)} \end{aligned} \quad . \tag{8d}$$

Noticing that this differs from (7d) of [1] only by the term $\prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu'_{j-1}}$ which is left unchanged in the process of combining the kernel of the prior density with the kernel of the likelihood, we may use the results of the algebra of the proof of (7) in [1] to give the posterior kernel as

$$e^{-\frac{1}{2}n''(\underline{\mu}-\underline{m}'')^t\underline{\underline{h}}(\underline{\mu}-\underline{m}'')} \ |\underline{\underline{h}}|^{\frac{1}{2}\delta''} \ e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}^*} \ |\underline{\underline{h}}|^{\frac{1}{2}\nu''_r-1} \ \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu'_{j-1}}$$

which is the kernel of an extended type 1 Normal-Wishart density with parameter $(\underline{m}'', \underline{\underline{V}}'', n'', \underline{\nu}'')$.

2.3 Mean Vector and Variance-Covariance Matrix of the Marginal Distribution of $\tilde{\underline{\mu}}$

When $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ has a extended type 1 Normal-Wishart distribution with parameter $(\underline{m}, \underline{\underline{V}}, \underline{\nu})$, then

$$E(\tilde{\underline{\mu}}) = \underline{m} \quad \text{and} \quad V(\tilde{\underline{\mu}}) = E(\tilde{\underline{\underline{h}}}^{-1}) \quad . \tag{9a}$$

Furthermore,

$$E(\tilde{\underline{\underline{h}}}^{-1}) = \underline{\underline{T}}^{-1} \underline{\underline{\eta}}\, \underline{\underline{T}}^{t^{-1}} \tag{9b}$$

where $\underline{\underline{T}}$ is an $(r \times r)$ upper triangular matrix such that $\underline{\underline{T}}^t\underline{\underline{T}}=\underline{\underline{V}}^{-1}$, and $\underline{\underline{\eta}}$ is an $(r \times r)$ diagonal matrix with first diagonal element $\eta_{11}=(a_r-2)^{-1}$, and for $j=2,3,\ldots,r$,

$$\eta_{jj} = \frac{[a_{r-j+2}-\frac{1}{2}(r-j+2)]}{[a_{r-j+1}-\frac{1}{2}(r-j+2)]}\, \eta_{j-1,j-1} \tag{9c}$$

Proof: That $E(\tilde{\underline{\mu}})=\underline{m}'$ follows from

$$E_{\underline{\underline{h}}}\, E_{\underline{\mu}|\underline{\underline{h}}}(\tilde{\underline{\mu}}) = E_{\underline{\underline{h}}}(\underline{m}') = \underline{m}' \quad .$$

That $V(\tilde{\underline{\mu}})=E(\tilde{\underline{\underline{h}}}^{-1})$ follows from formula (23) of [3]:

$$V(\tilde{\underline{\mu}}) = E_{\underline{\underline{h}}}\, V_{\underline{\mu}|\underline{\underline{h}}}(\tilde{\underline{\mu}}) + V_{\underline{\underline{h}}}\, E_{\underline{\mu}|\underline{\underline{h}}}(\tilde{\underline{\mu}}) \quad .$$

Since $E_{\underline{\mu}|\underline{\underline{h}}}(\tilde{\underline{\mu}})=\underline{m}'$, a constant not depending on $\underline{\underline{h}}$ and since $V_{\underline{\mu}|\underline{\underline{h}}}(\tilde{\underline{\mu}})=\underline{\underline{h}}^{-1}$, the first term above is $E(\tilde{\underline{\underline{h}}}^{-1})$ and the second is the null matrix $\underline{\underline{0}}$.

To prove (9b) and (9c), let $\underline{y}^{[j]}=\underline{\underline{T}}^{[j]}(\underline{\mu}-\underline{m})^{[j]}$ where $\underline{\underline{T}}$ is the upper triangular matrix such that $\underline{\underline{T}}^{-1}\underline{\underline{T}}^{t^{-1}}=\underline{\underline{V}}$, and make an integrand transform from $\underline{\mu}$ to $\underline{y}$ in (6a), giving as in (6e)

$$D(\underline{y}|\underline{\nu}) = S(\underline{\underline{I}}, \underline{\nu})\, (1 + \underline{y}^t\underline{y})^{-a_1-\frac{1}{2}} \prod_{j=1}^{r-1} (1 + \underline{y}^{[j]^t}\underline{y}^{[j]})^{-\nu_j} \quad .$$

Then

$$\eta_{11} \equiv V(\tilde{y}_1) = E(\tilde{y}_1^2) = \int_{R_{\underline{y}}} y_1^2\, D(\underline{y}|\underline{\nu})d\underline{y}$$

or letting $x_i=y_i^2$, $i=1,2,\ldots,r$, and making this integrand transform in the above integral,

$$V(\tilde{y}_1) = 2^{-r}\, S(\underline{\underline{I}}, \underline{\nu}) \int_0^\infty \cdots \int_0^\infty x_1^{\frac{1}{2}} \, [\prod_{j=2}^{r} x_j^{-\frac{1}{2}}]$$

$$\cdot\, (1 + \sum_{k=1}^{r} x_k)^{-a_1-\frac{1}{2}} [\prod_{j=1}^{r-1} (1+\sum_{\ell=1}^{j} x_\ell)^{-\nu_j}]dx_1\ldots dx_r$$

$$= S(\underline{\underline{I}},\ \underline{\nu})\ [2^{-r}\ B(\tfrac{1}{2},a_1)\ B(\tfrac{1}{2},a_2-\tfrac{1}{2})\ldots B(\tfrac{1}{2},a_{r-1}-\tfrac{1}{2}(r-2))\ B(\tfrac{1}{2},a_r-\tfrac{1}{2}(r+1))]$$

$$= \frac{1}{2(a_r-2)} = \frac{1}{\nu_r-2} \quad .$$

Evaluating $E(\tilde{y}_i^2)$ for $i=2,3,\ldots,r$ in a similar fashion yields (9c). Then (9b) follows directly.

2.4 Mean Matrix of $\underline{\underline{\tilde{h}}}^{-1} = \underline{\underline{\tilde{\sigma}}}$ and Variance of $\tilde{\sigma}_{ii}$

We have $E(\underline{\underline{\tilde{h}}}^{-1}) = V(\underline{\tilde{\mu}})$ in (9b). Here we derive the variance of $\tilde{\sigma}_{\alpha\alpha}$, $\alpha=1,2,\ldots,r$ as a function of the expected values of products $\tilde{y}_i y_j \tilde{y}_k \tilde{y}_\ell$ of elements of $\underline{\tilde{y}}$ as defined in the proof of (9c), and explicitly evaluate $V(\tilde{\sigma}_{11})$ and $V(\tilde{\sigma}_{22})$ for $r=2$. As mentioned earlier, we may use the ideas of the proof of (9c) to evaluate all $V(\tilde{\sigma}_{\alpha\beta})$ and $\mathrm{cov}(\tilde{\sigma}_{\alpha\beta},\ \tilde{\sigma}_{\gamma\delta})$ terms, and will do so in a later draft.

We now show that if an extended Type 1 Normal-Wishart distribution with parameter $(\underline{m},\ n,\ \underline{\underline{V}},\ \underline{\nu})$ is assigned to $(\underline{\tilde{\mu}},\ \underline{\underline{\tilde{h}}})$,

$$V(\tilde{h}^{-1}_{\alpha\alpha}) = V(\tilde{\sigma}_{\alpha\alpha}) = \frac{1}{3} \sum_{i,j,k,}^{r} t^{-1}_{\alpha i}\ t^{-1}_{\alpha j}\ t^{-1}_{\alpha k}\ t^{-1}_{\alpha \ell}\ E(\tilde{y}_i \tilde{y}_j \tilde{y}_k \tilde{y}_\ell) - E^2(\tilde{\sigma}_{\alpha\alpha}) \quad (10a)$$

where $t^{-1}_{\gamma\delta}$ is the $\gamma\delta$th element of the upper triangular matrix $\underline{\underline{T}}$, $\underline{\underline{T}}^t\underline{\underline{T}} = \underline{\underline{V}}^{-1}$, the sum is over all $1 \leq i,j,j,\ell \leq r$ and the $\tilde{y}_i$s are elements of $\underline{\tilde{y}}$, distributed as in (6e).

Proof: From formula (23) in Martin [3], $E([\tilde{h}^{-1}_{\alpha\alpha}]^2) = E(\tilde{\sigma}^2_{\alpha\alpha}) = \frac{1}{3} E(\tilde{\mu}^4_\alpha)$. As the matrix $\underline{\underline{T}}$ is such that $\underline{\underline{T}}^{-1}\underline{y} = \underline{\mu}$, $\mu_\alpha = \sum_{j=1}^{r} t^{-1}_{\alpha j}\ y_j$, $\alpha=1,2,\ldots,r$, and so

$$E(\tilde{\mu}^4_\alpha) = E([\sum_{j=1}^{r} t^{-1}_{\alpha j}\ \tilde{y}_j]^4) \quad ,$$

leading directly to (10a).

Notice that the density of $\underline{\tilde{y}}$ in (6e) is symmetric about $\underline{0}$ and the elements of $\underline{\tilde{y}}$ are uncorrelated. Hence only terms involving all even powers will be non-zero, i.e. $E(\tilde{y}_i^2\ \tilde{y}_j^2)$, $i \neq j$, and $E(\tilde{y}_j^4)$.

When r=2,

$$V(\tilde{\sigma}_{11}) = \frac{\eta_{11}^2 (t_{11}^{-1})^4}{(a_2 - \frac{5}{2})} = \frac{E^2(\tilde{\sigma}_{11})}{(a_2 - \frac{5}{2})} \qquad (10b)$$

and

$$V(\tilde{\sigma}_{22}) = \frac{\eta_{11}}{2(a_2 - \frac{5}{2})} \{(t_{21}^{-1})^4 + 2(t_{21}^{-1})^2 (t_{22}^{-1})^2 \frac{\eta_{22}}{\eta_{11}} + (t_{22}^{-1})^2 [\frac{\eta_{22}}{\eta_{11}} \cdot \frac{a_2 - 2}{a_1 - 2}]\}$$
$$- E^2(\tilde{\sigma}_{22}) \quad . \ (10c)$$

Proof: From (10a) we have

$$E(\tilde{\mu}_1^4) = E([\sum_{j=1}^{2} t_{1j}^{-1} \tilde{y}_j]^4) = t_{11}^{-4} E(\tilde{y}_1^4)$$

and

$$E(\tilde{\mu}_2^4) = E([\sum_{j=1}^{2} t_{2j}^{-1} y_j]^4)$$

$$= (t_{21}^{-1})^4 E(\tilde{y}_1^4) + 2(t_{21}^{-1})^2 (t_{22}^{-1})^2 E(\tilde{y}_1 \tilde{y}_2) + (t_{22}^{-1})^4 E(\tilde{y}_2^4) \quad .$$

We may evaluate $E(\tilde{y}_i^4)$, i=1,2, and $E(\tilde{y}_1^2 \tilde{y}_2^2)$ directly, using the same ideas of the proof of (9b) to find

$$E(\tilde{y}_1^4) = \frac{3}{4(a_2 - \frac{3}{2})(a_2 - \frac{5}{2})} \quad ,$$

$$E(\tilde{y}_2^4) = \frac{3(a_2 - 1)(a_2 - 2)}{4(a_1 - 1)(a_1 - 2)(a_2 - \frac{3}{2})(a_2 - \frac{5}{2})} \quad ,$$

$$E(\tilde{y}_1 \tilde{y}_2) = \frac{(a_2 - 1)}{4(a_1 - 1)(a_2 - \frac{3}{2})(a_2 - \frac{5}{2})} \quad .$$

Using the fact that

$$V(\tilde{\sigma}_{ii}) = \frac{1}{3} E(\tilde{\mu}_i^4) - E^2(\tilde{\sigma}_{ii}) \quad , \qquad i=1,2,$$

the above, and (9b), we have (10b) and (10c).

2.5 An Example

To illustrate how the results of section 2.4 give us more latitude in assigning a prior to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ we treat a (2 x 2) numerical example, first using a Normal-Wishart prior and then using an extended Type 1 Normal-Wishart prior. In each case we wish to assign a prior such that

$$E(\tilde{\underline{\mu}}) = \begin{pmatrix} 200 \\ 200 \end{pmatrix} , \quad V(\tilde{\underline{\mu}}) = \begin{bmatrix} 11.0 & 5.5 \\ 5.5 & 11.0 \end{bmatrix} , \quad E(\tilde{\underline{\underline{\sigma}}}) = \begin{bmatrix} 110 & 55 \\ 55 & 110 \end{bmatrix} ,$$

$$V(\tilde{\sigma}_{11}) = 605.0 , \qquad V(\tilde{\sigma}_{22}) = 1210. \qquad .$$

After observing a sample yielding statistics

$$n = 10 , \quad \nu = 8 , \quad \underline{m} = \begin{pmatrix} 150.0 \\ 250.0 \end{pmatrix} , \quad \underline{\underline{V}} = \begin{bmatrix} 1360. & 120.0 \\ 120.0 & 330.0 \end{bmatrix} ,$$

$$\frac{1}{\nu-2} \underline{\underline{V}} = \begin{bmatrix} 220 & 20 \\ 20 & 55 \end{bmatrix} ,$$

we also wish to find the posterior of $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ in both cases.

2.5.1 Analysis Using Normal-Wishart Prior

We may match $E(\tilde{\underline{\mu}})$, $E(\tilde{\underline{\underline{\sigma}}})$, and $V(\tilde{\underline{\mu}})$ by setting

$$E(\tilde{\underline{\mu}}) = \underline{m}' = \begin{pmatrix} 200 \\ 200 \end{pmatrix} , \quad E(\tilde{\underline{\underline{\sigma}}}) = \frac{1}{\nu'-2} \underline{\underline{V}}' = \begin{bmatrix} 110 & 55 \\ 55 & 110 \end{bmatrix} ,$$

$$V(\tilde{\underline{\mu}}) = \frac{1}{n'} E(\tilde{\underline{\underline{\sigma}}}) = \begin{bmatrix} 11.0 & 5.5 \\ 5.5 & 11.0 \end{bmatrix}$$

so that $n'=10$. We must, however use $V(\tilde{\sigma}_{11})$ and $V(\tilde{\sigma}_{22})$ to determine ν' and $\underline{\underline{V}}'$, and as

$$V(\tilde{\sigma}_{11}) = \frac{2E^2(\tilde{\sigma}_{11})}{\nu'-4} = \frac{24,200}{\nu'-4} \quad \text{and} \quad V(\tilde{\sigma}_{22}) = \frac{2E^2(\tilde{\sigma}_{22})}{\nu'-4} = \frac{24,200}{\nu'-4}$$

when a Normal-Wishart prior is assigned to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$, it is impossible to satisfy both prior specifications $V(\tilde{\sigma}_{11}) = 605.0$ and $V(\tilde{\sigma}_{22}) = 1210.0$. We arbitrarily choose the former, yielding

$$\nu' = 44 \quad , \quad \underline{\underline{V}}' = \begin{bmatrix} 4620 & 2310 \\ 2310 & 4620 \end{bmatrix} .$$

Once the sample has been observed formula (7) of [1] gives posterior parameters

$$n'' = n' + n = 10 + 10 = 20 ,$$

$$\nu'' = \nu' + \nu + r + \delta + \delta' - \delta'' - \Phi - 1 = 44 + 10 - 1 = 53 ,$$

$$n_u = n'n/n'' = (10)(10)/20 = 5 \quad , \quad \text{(redundant)} ,$$

$$\underline{\underline{V}}'' = V' + V + n_u(\underline{m}' - \underline{m})(\underline{m}' - \underline{m})^t = \begin{bmatrix} 18,480 & -10,070 \\ -10,070 & 17,450 \end{bmatrix} ,$$

$$\underline{m}'' = n''^{-1}(n'\underline{m}' + n\,\underline{m}) = \begin{pmatrix} 175 \\ 225 \end{pmatrix} .$$

Posterior to the sample we have

$$E(\tilde{\underline{\mu}}) = \begin{pmatrix} 175 \\ 225 \end{pmatrix} , \quad E(\tilde{\underline{\underline{\sigma}}}) = \frac{1}{\nu''-2}\,\underline{\underline{V}}'' = \begin{bmatrix} 362.4 & -197.5 \\ -197.5 & 342.2 \end{bmatrix} ,$$

$$V(\tilde{\underline{\mu}}) = \frac{1}{n''}\,E(\tilde{\underline{\underline{\sigma}}}) = \begin{bmatrix} 18.12 & -9.875 \\ -9.875 & 17.11 \end{bmatrix} ,$$

and

$$V(\tilde{\sigma}_{11}) = \frac{2E^2(\tilde{\sigma}_{11})}{\nu''-4} = 5361,$$

$$V(\tilde{\sigma}_{22}) = \frac{2E^2(\tilde{\sigma}_{22})}{(\nu''-4)} = 4780. $$

Notice that the sample has drastically increased the variances of $\tilde{\underline{\mu}}$, of $\tilde{\sigma}_{11}$ and of $\tilde{\sigma}_{22}$. This occurs because the sample mean vector is very far from $\underline{m}'$ when measured in units of the standard deviation of $\tilde{\mu}_1$ and $\tilde{\mu}_2$ as specified in the prior.

2.5.2 Analysis Using Extended Type 1 Normal-Wishart Prior

If we use the extended natural conjugate prior family, we may match $E(\tilde{\underline{\mu}})$, $E(\tilde{\underline{\underline{\sigma}}})$, $V(\tilde{\underline{\mu}})$, $V(\tilde{\sigma}_{11})$, and $V(\tilde{\sigma}_{22})$. To find the parameters $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}')$ in this case, we note that

$$E(\tilde{\underline{\mu}}) = \underline{m}' = \begin{pmatrix} 200 \\ 200 \end{pmatrix} ,$$

$$E(\tilde{\underline{\underline{\sigma}}}) = \underline{\underline{T}}'^{-1} \underline{\underline{\eta}} \, \underline{\underline{T}}'^{t^{-1}} = \begin{bmatrix} 110 & 55 \\ 55 & 110 \end{bmatrix}$$

$$V(\tilde{\underline{\mu}}) = n' E(\tilde{\underline{\underline{\sigma}}}) = \begin{bmatrix} 11.0 & 5.5 \\ 5.5 & 11.0 \end{bmatrix}$$

and

$$V(\tilde{\sigma}_{11}) = \frac{2E^2(\tilde{\sigma}_{11})}{\nu_2'-4} = 605.0$$

$$V(\tilde{\sigma}_{22}) = \frac{1}{4(a_2'-\frac{3}{2})(a_2'-\frac{5}{2})} \{(t_{21}'^{-1})^4 + 2(t_{21}'^{-1})^2 (t_{22}'^{-1})^2 \frac{\eta_{22}'}{\eta_{11}'}$$

$$+ (t_{22}'^{-1})^4 \frac{\eta_{22}'}{\eta_{11}'} \cdot \frac{a_2'-2}{a_1'-2}\} - E^2(\tilde{\sigma}_{22}) = 1210. \quad .$$

Solving the above formulae yields

$$\underline{m}' = \begin{pmatrix} 200 \\ 200 \end{pmatrix} , \qquad n' = 10 \quad , \qquad \underline{\nu}' = (44, 13.23)$$

$$\underline{\underline{T}}'^{-1} = \begin{bmatrix} 67.97 & 0 \\ 33.99 & 36.51 \end{bmatrix} , \qquad \underline{\underline{\eta}}' = \begin{bmatrix} .02381 & 0 \\ 0 & .06179 \end{bmatrix} ,$$

$$\underline{\underline{V}}' = \underline{\underline{T}}'^{-1} \, \underline{\underline{T}}'^{t^{-1}} = \begin{bmatrix} 4620 & 2310 \\ 2310 & 2488 \end{bmatrix} .$$

After observing the sample, we may compute the posterior parameters:

$$n'' = n' + n = 10 + 10 = 20 ,$$

$$\nu''_2 = \nu'_2 + \nu + r - \delta + \delta' - \delta'' - \Phi - 1 = 44 + 10 - 1 = 53 ,$$

$$\nu''_1 = \nu'_1 = 13.23 ,$$

$$n_u = n'n/n'' = (10)(10)/20 = 5 , \quad \text{(redundant)} ,$$

$$\underline{\underline{V}}'' = \underline{\underline{V}}' + \underline{\underline{V}} + n_u(\underline{m}' - \underline{m})(\underline{m}' - \underline{m})^t = \begin{bmatrix} 18{,}480 & -10{,}070 \\ -10{,}070 & 15{,}318 \end{bmatrix} ,$$

$$\underline{m}'' = n''^{-1}(n'\underline{m}' + n\underline{m}) = \begin{pmatrix} 175 \\ 225 \end{pmatrix} .$$

From these parameters we may compute

$$\underline{\underline{T}}''^{-1} = \begin{bmatrix} 135.8 & 0 \\ -74.15 & 99.10 \end{bmatrix} , \qquad \underline{\underline{\eta}}'' = \begin{bmatrix} .01961 & 0 \\ 0 & .03992 \end{bmatrix}$$

leading to

$$E(\underline{\underline{\tilde{\sigma}}}) = \underline{\underline{T}}''^{-1} \, \underline{\underline{\eta}}'' \, \underline{\underline{T}}''^{t^{-1}} = \begin{bmatrix} 362.4 & -197.5 \\ -197.5 & 499.8 \end{bmatrix} ,$$

$$V(\underline{\tilde{\mu}}) = \frac{1}{n''} E(\underline{\underline{\tilde{\sigma}}}) = \begin{bmatrix} 181.2 & -98.75 \\ -98.75 & 249.9 \end{bmatrix} ,$$

$$V(\tilde{\sigma}_{11}) = \frac{E^2(\tilde{\sigma}_{11})}{a''_2 - \frac{5}{2}} = 5361. \quad ,$$

and

$$V(\tilde{\sigma}_{22}) = \frac{1}{4(a''_2-\frac{3}{2})(a''_2-\frac{5}{2})} \{\cdot\} - E^2(\tilde{\sigma}_{22}) = 82,227 \quad .$$

In this case, the posterior variance of $\tilde{\sigma}_{22}$ is drastically larger than the prior variance of $\tilde{\sigma}_{22}$ and also much larger than that of $\tilde{\sigma}_{22}$ when a Normal-Wishart prior was assigned.

3. Sampling Distributions with Fixed n

We assume that a sample of size n is to be drawn from an r-dimensional independent Multinormal process whose parameter $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ is a random variable having an extended Type 1 Normal-Wishart distribution with parameter $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}')$.

3.1 Conditional Joint Distribution of $(\tilde{\underline{m}}, \tilde{\underline{\underline{V}}} | \underline{\mu}, \underline{\underline{h}})$

The conditional joint distribution of the statistic $(\tilde{\underline{m}}, \tilde{\underline{\underline{V}}})$ defined in formula (8) of [1] given that the process parameter has value $(\underline{\mu}, \underline{\underline{h}})$ is, provided $\nu > 0$,

$$D(\underline{m}, \underline{\underline{V}} | \underline{\mu}, \underline{\underline{h}}, \nu) = f_N^{(r)}(\underline{m} | \underline{\mu}, \underline{\underline{h}}n)\, f_W^{(r)}(\underline{\underline{V}} | \underline{\underline{h}}, \nu)$$

3.2 Unconditional Joint Distribution of $(\tilde{\underline{m}}, \tilde{\underline{\underline{V}}})$

The unconditional (with respect to $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$) joint distribution of $(\tilde{\underline{m}}, \tilde{\underline{\underline{V}}})$ when $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ has been assigned an extended Type 1 Normal-Wishart prior with parameter $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}')$ has density

$$\begin{aligned} &D(\underline{m}, \underline{\underline{V}} | \underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu) \\ &= \int_{R_{\underline{\mu}}} \int_{R_{\underline{\underline{h}}}} f_N^{(r)}(\underline{m} | \underline{\mu}, \underline{\underline{h}}n)\, f_W^{(r)}(\underline{\underline{V}} | \underline{\underline{h}}, \nu)\, f_{NW1}^{(r)}(\underline{\mu}, \underline{\underline{h}} | \underline{m}', \underline{\underline{V}}', n', \underline{\nu}')\, d\underline{\mu}\, d\underline{\underline{h}} \end{aligned} \tag{11a}$$

where the domain of integration of $\underline{\mu}$ is $R_{\underline{\mu}} = (-\underline{\infty}, +\underline{\infty})$ and of $\underline{\underline{h}}$ is $R_{\underline{\underline{h}}}$.

It follows that if $\nu_r' > 0$, $n' > 0$, $\underline{\underline{V}}'$ is PDS, and $a_j' > \frac{1}{2}(j-1)$,

$$D(\underline{m}, \underline{\underline{V}} | \underline{m}', \underline{\underline{V}}', n', \nu'; n, \nu) \propto |\underline{\underline{V}}|^{\frac{1}{2}\nu - 1}\, |\underline{\underline{B}}|^{\beta' - \frac{1}{2}(\nu'' + r - 1)} \cdot \prod_{j=1}^{r-1} |\underline{\underline{B}}[j]|^{-\nu_j'} \tag{11b}$$

where

$$\underline{\underline{B}} \equiv n_u(\underline{m} - \underline{m}')(\underline{m} - \underline{m}')^t + \underline{\underline{V}}' + \underline{\underline{V}} \quad , \qquad \beta' = \sum_{j=1}^{r-1} \nu_j' \ .$$

Proof: Following the initial steps in the proof of formula (10) in [1], we may write, letting $n_u = n'n/n''$,

$$D(\underline{m},\ \underline{\underline{V}}|\underline{m}',\ \underline{\underline{V}}',\ n',\ \underline{\nu}';\ n,\ \nu)$$

$$= \int_{\underline{\underline{h}}} f_N^{(r)}(\underline{m}|\underline{m}',\ \underline{\underline{h}}n_u)\ f_W^{(r)}(\underline{\underline{V}}|\underline{\underline{h}},\ \nu)\ f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}',\ \underline{\nu}')d\underline{\underline{h}} \quad .$$

This integral is proportional to

$$|\underline{\underline{V}}|^{\frac{1}{2}\nu-1} \int_{\underline{\underline{h}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\{n_u(\underline{m}-\underline{m}')(\underline{m}-\underline{m}')^t+\underline{\underline{V}}+\underline{\underline{V}}'\}}\ |\underline{\underline{h}}|^{\frac{1}{2}\nu''_{r}-1} \cdot \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu'_{j-1}}\ d\underline{\underline{h}}$$

Letting

$$\underline{\underline{B}} = n_u(\underline{m}-\underline{m}')(\underline{m}-\underline{m}')^t + \underline{\underline{V}} + \underline{\underline{V}}'$$

the above integral may be written as

$$|\underline{\underline{V}}|^{\frac{1}{2}\nu-1} \int_{\underline{\underline{h}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{B}}}\ |\underline{\underline{h}}|^{\frac{1}{2}\nu''_{r}-1}\ \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu'_{j-1}}\ d\underline{\underline{h}} \quad .$$

The integrand is the kernel of an extended Type 1 Wishart density with parameter $(\underline{\underline{B}},\ \underline{\nu}'')$ where $\underline{\nu}''=(\nu'_1,\ldots,\nu'_{r-1},\ \nu''_r)$. Hence apart from a normalizing constant depending on neither $\underline{\underline{V}}$ nor $\underline{\underline{B}}$,

$$D(\underline{m},\underline{\underline{V}}|\underline{m}',\underline{\underline{V}}',n',\nu';n,\nu) \propto |\underline{\underline{V}}|^{\frac{1}{2}\nu-1}\ |\underline{\underline{B}}|^{\beta'-\frac{1}{2}(\nu''_r+r-1)}\ \prod_{j=2}^{r-1} |\underline{\underline{B}}^{[j]}|^{-\nu_j} \quad .$$

3.3 Unconditional Distribution of $\underline{\tilde{m}}$

When $n_u > 0$, $a'_j > \frac{1}{2}(j-1)$, $\nu'_r > 0$, and $\underline{\underline{V}}'$ is PDS,

$$D(\underline{m}|\underline{m}',\ \underline{\underline{V}}',\ n',\ \underline{\nu}';\ n,\ \nu) \qquad \text{(12a)}$$

$$\propto\ (1 + n_u(\underline{m}-\underline{m}')^t\underline{\underline{V}}'^{-1}(\underline{m}-\underline{m}))^{\beta'-\frac{1}{2}(\nu'_r+r)} \cdot \prod_{j=1}^{r-1} (1+n_u(\underline{m}-\underline{m}')^{[j]^t}\underline{\underline{V}}^{[j]^{-1}}(\underline{m}-\underline{m}')^{[j]})^{-\nu'_j}$$

so that if, in addition, $2\nu'_j-j > 0$, $j=2,3,\ldots,r-1$,

$$D(\underline{m}|\underline{m}', \underline{\underline{V}}', \underline{\nu}'; n, \nu)$$

$$f_S^{(r)}(\underline{m}|\underline{m}', \underline{\underline{H}}_\nu, \nu_r') \prod_{j=1}^{r-1} f_S^{(j)}(\underline{m}^{[j]}|\underline{m}'^{[j]}, \underline{\underline{H}}_j, 2\nu_j'-j) \tag{12b}$$

where $\underline{\underline{H}}_\nu \equiv (\nu_r'-2\beta')n_u \underline{\underline{V}}'^{-1}$ and $\underline{\underline{H}}_j = (2\nu_j'-j)n_u \underline{\underline{V}}'^{[j]^{-1}}$, $j=1,2,\ldots,r-1$. Both (12a) and (12b) hold even if $\nu \leq 0$.

Proof: The kernel of the marginal likelihood (4b') in [1] of $\underline{\tilde{m}}$ given the parameter $(\underline{\mu}, \underline{\underline{h}})$ and that $n > 0$ is Normal with parameter $(\underline{\mu}, \underline{\underline{h}}n)$. Also, the kernel of the prior of $\underline{\tilde{\mu}}$ given $\underline{\underline{\tilde{h}}}=\underline{\underline{h}}$ is Normal with parameter $(\underline{m}, \underline{\underline{h}}n')$.

Since $\underline{\tilde{\mu}}$ and $\underline{\tilde{\epsilon}} \equiv \underline{\tilde{m}}-\underline{\tilde{\mu}}$ are, given $\underline{\underline{\tilde{h}}}=\underline{\underline{h}}$, each independent Normal random vectors and since a sum of independent (r x 1) Normal random vectors is (r x 1) Normal, it follows that $\underline{\tilde{m}}=\underline{\tilde{\mu}}+\underline{\tilde{\epsilon}}$ is Normal with mean vector

$$E(\underline{\tilde{m}}) = E(\underline{\tilde{\mu}}) + E(\underline{\tilde{\epsilon}}) = E(\underline{\tilde{\mu}}) = \underline{m}'$$

and variance-covariance matrix

$$V(\underline{\tilde{m}}) = V(\underline{\tilde{\mu}}) + V(\underline{\tilde{\epsilon}}) = \underline{\underline{h}}^{-1}\left(\frac{1}{n'}+\frac{1}{n}\right) = (\underline{\underline{h}}n_u)^{-1} \quad .$$

Hence the distribution of $\underline{\tilde{m}}$ unconditional as regards $\underline{\tilde{\mu}}$ but conditional upon $\underline{\underline{\tilde{h}}}=\underline{\underline{h}}$ is Normal with parameter $(\underline{m}', \underline{\underline{h}}n_u)$.

To find the distribution of $\underline{\tilde{m}}$ unconditional as regards $\underline{\underline{\tilde{h}}}$ as well as $\underline{\tilde{\mu}}$, we note that the marginal distribution of $\underline{\underline{\tilde{h}}}$ is from (7) $f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}', \underline{\nu}')$ so

$$D(\underline{m}|\underline{m}',\underline{\underline{V}}',n',\underline{\nu}';n,\nu) = \int_{R_{\underline{\underline{h}}}} f_N^{(r)}(\underline{m}|\underline{m}', \underline{\underline{h}}n_u)\, f_{W1}^{(r)}(\underline{\underline{h}}|\underline{\underline{V}}', \underline{\nu}')d\underline{\underline{h}} \quad .$$

Since $n > 0$ and $n' > 0$ here, $\delta=1$ and the above is proportional to

$$\int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}n_u(\underline{m}-\underline{m}')^t\underline{\underline{h}}(\underline{m}-\underline{m}')} |\underline{\underline{h}}|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}'} |\underline{\underline{h}}|^{\frac{1}{2}\nu_r'-1} \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu_j'-1} d\underline{\underline{h}}$$

$$= \int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\{n_u(\underline{m}-\underline{m}')(\underline{m}-\underline{m}')^t+\underline{\underline{V}}'\}} |\underline{\underline{h}}|^{\frac{1}{2}(\nu_r'-1)} \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu_j'-1} d\underline{\underline{h}} \quad .$$

Observing that the integrand in the integral immediately above is the kernel of an extended Type 1 Wishart distribution with parameter $(n_u(\underline{m}-\underline{m}')(\underline{m}-\underline{m}')^t+\underline{\underline{V}}', \underline{\nu}_o)$ where $\underline{\nu}_o=(\nu_1',\ldots,\nu_{r-1}',\nu_r'+1)$. Both (12a) and (12b) follow directly.

3.4 Unconditional Distribution of $\tilde{\underline{\underline{V}}}$ When $\nu > 0$

When $\nu > 0$, $\nu' > 0$, and $\underline{\underline{V}}'$ is PDS, the kernel of the distribution of $\tilde{\underline{\underline{V}}}$ unconditional as regards $\tilde{\underline{m}}$, $\tilde{\underline{\mu}}$, and $\tilde{\underline{\underline{h}}}$ is

$$D(\underline{\underline{V}}|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu) \tag{13}$$
$$\propto \frac{|\underline{\underline{V}}|^{\frac{1}{2}\nu-1}}{|\underline{\underline{V}}+\underline{\underline{V}}'|^{-\beta'+\frac{1}{2}(\nu_r''+r)-1}} \cdot \prod_{j=1}^{r-1} |[\underline{\underline{V}} + \underline{\underline{V}}']^{[j]}|^{-\nu_j'} \quad .$$

Proof: From (4c) of [1] the marginal likelihood of $\tilde{\underline{\underline{V}}}$ is proportional to

$$|\underline{\underline{V}}|^{\frac{1}{2}\nu-1} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}\ \underline{\underline{V}}} |\underline{\underline{h}}|^{\frac{1}{2}(\nu+r-1)} \quad .$$

Multiplying this kernel by the kernel of the prior distribution (5) of $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ with parameter $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}')$ and integrating over the range of $\tilde{\underline{\mu}}$ gives, when $n' > 0$,

$$D(\underline{\underline{V}}, \underline{\underline{h}}|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$$
$$\propto |\underline{\underline{V}}|^{\frac{1}{2}\nu-1} |\underline{\underline{h}}|^{\frac{1}{2}(\nu_r''-1)-1} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}(\underline{\underline{V}}+\underline{\underline{V}}')} \quad .$$

Integrating the above kernel over $R_{\underline{\underline{h}}}$, we have

$$D(\underline{\underline{V}}|\underline{m}', \underline{\underline{V}}', n', \nu'; n, \nu)$$
$$\propto |\underline{\underline{V}}|^{\frac{1}{2}\nu-1} \int_{R_{\underline{\underline{h}}}} e^{-\frac{1}{2}\mathrm{tr}\ \underline{\underline{h}}(\underline{\underline{V}}+\underline{\underline{V}}')} |\underline{\underline{h}}|^{\frac{1}{2}(\nu_r''-1)-1} \prod_{j=2}^{r} |\underline{\underline{h}}_{[j]}|^{-\nu_{j-1}'} d\underline{\underline{h}} \quad .$$

The integrand in the above integral is the kernel of an extended Type 1 Wishart density with parameter $(\underline{\underline{V}}+\underline{\underline{V}}', \underline{\nu}_o)$ where $\underline{\nu}_o=(\nu_1',\ldots,\nu_{r-1}'), \nu_r''-1)$. Formula (13) follows directly.

4. Preposterior Analysis with Fixed $n > 0$

We assume that a sample of fixed size $n > 0$ is to be drawn from an r-dimensional Multinormal process with mean vector $\underline{\mu}$ and matrix precision $\underline{\underline{h}}$ are not known with certainty but are regarded as random variables $(\tilde{\underline{\mu}}, \tilde{\underline{\underline{h}}})$ having a extended Type 1 Normal-Wishart prior distribution with parameter $(\underline{m}, \underline{\underline{V}}', n', \underline{\nu}')$ where $n' > 0$, $\nu'_r > 0$, $a'_j > \frac{1}{2}(j-1)$; however, $\underline{\underline{V}}'$ may or may not be singular.

4.1 Joint Distribution of $(\underline{m}'', \underline{\underline{V}}'')$

The joint density of $(\underline{m}'', \underline{\underline{V}}'')$ provided $n' > 0$, $\nu'_r > 0$, $\nu > 0$, and $a'_j > \frac{1}{2}(j-1)$ is,

$$D(\underline{m}'', \underline{\underline{V}}''|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$$

$$\propto \frac{|\underline{\underline{V}}''-\underline{\underline{V}}'-n^*(\underline{m}''-\underline{m}')(\underline{m}''-\underline{m}')^t|^{\frac{1}{2}\nu-1}}{|\underline{\underline{V}}''|^{-\beta'+\frac{1}{2}(\nu''+r-1)}} \cdot \prod_{j=1}^{r-1} |\underline{\underline{V}}''[j]|^{-\nu'_j} \tag{14a}$$

where

$$n^* = n'n''/n \tag{14b}$$

and the range of $(\underline{m}'', \underline{\underline{V}}'')$ is

$$R_{(\underline{m}'',\underline{\underline{V}}'')} \equiv \{(\underline{m}'',\underline{\underline{V}}'')|-\underline{\infty} < \underline{m}'' < +\underline{\infty}, \text{ and } \underline{\underline{V}}''-\underline{\underline{V}}'-n^*(\underline{m}''-\underline{m}')(\underline{m}''-\underline{m}')^t \text{ is PDS}\} .$$

Proof: From formulas (7), (18a) and (18b) of [1], we have

$$(\underline{m}, \underline{\underline{V}}) = (\tfrac{1}{n}(n''\underline{m}''-n'\underline{m}') , \underline{\underline{V}}''-\underline{\underline{V}}'-n^*(\underline{m}''-\underline{m}')(\underline{m}''-\underline{m}')^t) , \tag{15}$$

and from the proof of (17) in [1], the fact that $J(\underline{m}'', \underline{\underline{V}}''; \underline{m}, \underline{\underline{V}})$, the Jacobian of the integrand transformation from $(\underline{m}, \underline{\underline{V}})$ to $(\underline{m}'', \underline{\underline{V}}'')$, is a constant involving neither $\underline{m}$, nor $\underline{\underline{V}}$.

Making the transformation (15) in (11), since the Jacobian is just a constant of proportionality, we obtain (14a).

As in (17a) of [1], this density exists only if $\nu > 0$, since the numerator in the first term above vanishes when $\nu < 0$. The kernel exists, however even if $\underline{\underline{V}}'$ is singular provided that $\nu > 0$.

4.2 Some Distribution of $\underline{\tilde{m}}''$

The distribution of $\underline{\tilde{m}}''$ unconditional as regards $\underline{\tilde{\mu}}$ and $\underline{\underline{\tilde{h}}}$ provided $n > 0$, $\nu'_r > 0$, $a'_j > \frac{1}{2}(j-1)$, $j=1,2,\ldots,r-1$, and $\underline{\underline{V}}'$ is PDS, is

$$D(\underline{m}''|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$$

$$\propto (\nu'_r-2\beta'+(\underline{m}''-\underline{m}')^t\underline{\underline{H}}^o_\nu(\underline{m}''-\underline{m}'))^{\beta'-\frac{1}{2}(\nu'_r+r)} \tag{16a}$$

$$\cdot \prod_{j=1}^{r-1} (2\nu'_j-j)+(\underline{m}''^{[j]}-\underline{m}'^{[j]})^t\underline{\underline{H}}^o_j(\underline{m}''^{[j]}-\underline{m}'^{[j]}))^{-\nu'_j}$$

where

$$\underline{\underline{H}}^o_\nu = (\frac{n''}{n})^2 \underline{\underline{H}}_\nu = \frac{n''n'}{n} (\nu'_r-2\beta')\underline{\underline{V}}^{-1} \quad \text{and} \quad \underline{\underline{H}}^o_j = \frac{n''n'}{n} (2\nu'_j-j)\underline{\underline{V}}'^{[j]^{-1}} \quad .$$

If in addition, $2\nu'_j-j \geq 0$, $j=1,2,\ldots,r-1$, and $\nu'_r-2\beta' > 0$, we may write

$$D(\underline{m}''|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$$

$$\propto f^{(r)}_S(\underline{m}''|\underline{m}', \underline{\underline{H}}^o_\nu, \nu'_r-2\beta') \prod_{j=1}^{r-1} f^{(j)}_S(\underline{m}''^{[j]}|\underline{m}'^{[j]}, \underline{\underline{H}}^o_j, 2\nu'_j-j) \quad . \tag{16b}$$

Proof: The distribution of $\underline{\tilde{m}}$ unconditional as regards $\underline{\tilde{\mu}}$ and $\underline{\underline{\tilde{h}}}$ when $(\underline{\tilde{\mu}}, \underline{\underline{\tilde{h}}})$ is assigned an extended Type 1 Normal-Wishart prior is of the form (12) with parameter $(\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$. By Theorem 1 proven in substitution 2.2, since

$$\underline{\tilde{m}}'' = \frac{1}{n''} (n'\underline{m}' + n\underline{\tilde{m}}) \quad ,$$

the distribution of $\underline{\tilde{m}}''$ is of the form (12) with parameter $(\underline{m}', (\frac{n}{n''})^2\underline{\underline{V}}', n', \underline{\nu}'; n, \nu)$. Formula (16) follows by substitution of $(\frac{n}{n''})^2\underline{\underline{V}}'$ for $\underline{\underline{V}}'$ in (12).

The conditional distribution of $\underline{\tilde{m}}''$ given $\underline{\underline{\tilde{V}}}''=\underline{\underline{V}}''$ provided $\nu > 0$, $n' > 0$, $\nu'_r > 0$, and $a'_j > \frac{1}{2}(j-1)$, $j=1,2,\ldots,r-1$, is

$$D(\underline{m}''|\underline{m}', \underline{\underline{V}}', n', \underline{\nu}'; n, \nu, \underline{\underline{V}}'') = f^{(r)}_{iS}(\underline{m}''|\underline{m}', \underline{\underline{H}}_*, \nu) \tag{16a}$$

where

$$\underline{\underline{H}}_* = \nu n^*(\underline{\underline{V}}''-\underline{\underline{V}}')^{-1} \quad . \tag{16b}$$

Proof: The kernel of the conditional distribution of $\tilde{\underline{m}}''$ given $\tilde{\underline{\underline{V}}}''=\underline{\underline{V}}''$ is proportional to (14a), so

$$D(\underline{m}''|\underline{m}', \underline{\underline{V}}', n', \underline{v}'; n, \nu, \underline{\underline{V}}'')$$

$$\propto |\underline{\underline{V}}''-\underline{\underline{V}}'-n^*(\underline{m}''-\underline{m}')(\underline{m}''-\underline{m}'')^t|^{\frac{1}{2}\nu-1} \quad .$$

Hence, as shown in the proof of (20) in [1], this distribution of $\tilde{\underline{m}}$ is inverted Student with parameter $(\underline{m}', \underline{\underline{H}}_*, \nu)$, and $(\underline{\underline{V}}''-\underline{\underline{V}}')^{-1}$ is PDS as long as $\nu > 0$.

REFERENCES

[1] Albert Ando and G.M. Kaufman, "Bayesian Analysis of the Independent Multinormal Process--Neither Mean Nor Precision Known, Part I," Working Paper No. 41-63, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge (Revised March, 1964).

[2] R. Bellman, "A Generalization of Some Integral Identities Due to Ingham and Siegel," Duke Mathematical Journal, Vol. , No. , 1956, pp. 571-577.

[3] J. Martin, "Multinormal Bayesian Analysis: Two Examples," Working Paper No. 85-64, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, May, 1964.

[4] T.J. Rothenberg, "A Bayesian Analysis of Simultaneous Equation Systems," Report 6315, Netherlands Econometric Institute May, 1963.

[5] A. Zellner, unpublished.

MAR 1
APR 5 '71
MAR 1

BASEMENT

Date Due

[redacted]

JAN 03 '77

OCT 27 '77

~~NOV 29 '78~~

JUN 26 '79

SEP 11 '79

~~JAN [illegible]~~

JUL 7 1986

FEB 14 1990

MAY [illegible]

Lib-26-67

MIT LIBRARIES
3 9080 003 868 400
55-64

MIT LIBRARIES
3 9080 003 899 249
56-64

MIT LIBRARIES
3 9080 003 868 327
57-64

1751830

HD28
.M414
Nos.55-64
[illegible]

M.I.T. Alfred P. Sloan
School of Management
Working Papers.

MIT LIBRARIES
3 9080 003 868 426
62-64

MIT LIBRARIES
3 9080 003 868 277
65-64

MIT LIBRARIES
3 9080 003 868 715
66-64

MIT LIBRARIES
3 9080 003 868 707
67-64

MIT LIBRARIES
3 9080 003 868 681
70-64

MIT LIBRARIES
3 9080 003 868 673
75-64

MIT LIBRARIES
3 9080 003 899 694
78-64

MIT LIBRARIES
3 9080 003 899 645
80-64

MIT LIBRARIES
3 9080 003 899 678
81-64

MIT LIBRARIES
3 9080 003 899 710
82-64